

EV368629680

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION FOR LETTERS PATENT

**Systems and Methods for Parallel Evaluation of
Multiple Queries**

Inventors:

David Wortendyke

Umesh Madan

Geary L. Eppley

ATTORNEY'S DOCKET NO. MS1-1825US

RELATED APPLICATION

The present application is related to U.S. Patent Application No. 09/_____, entitled "Systems and Methods for Updating a Filter Engine Opcode Tree" and filed by the Applications named herein on February 20, 2004.

TECHNICAL FIELD

The systems and methods described herein generally relate to query processing and, more particularly, to systems and methods for parallel evaluation of multiple queries.

BACKGROUND

Computing systems - i.e. devices capable of processing electronic data such as computers, telephones, Personal Digital Assistants (PDA), etc. - communicate with other computing systems by exchanging data messages according to a communications protocol that is recognizable by the systems. Such a system utilizes filter engines containing queries that are used to analyze messages that are sent and/or received by the system and to determine if and how the messages will be processed further.

A filter engine may also be called an "inverse query engine." Unlike a database, wherein an input query is tried against a collection of data records, an inverse query engine tries an input against a collection of queries. Each query includes one or more conditions, or rules, that must be satisfied by an input for the query to evaluate to true against the input.

A collection of queries usually takes the form of one or more filter tables that may contain hundreds or thousands of queries, and each query may contain

1 several conditions. Significant system resources (e.g., setting up query contexts,
2 allocating buffers, maintaining stacks, etc.) are required to process an input against
3 each query in the filter table(s) and, therefore, processing an input against
4 hundreds or thousands of queries can be quite expensive.

5 Queries included in a particular system may be somewhat similar since the
6 queries are used within the system to handle data in a like manner. As a result,
7 several queries may contain common expressions that currently must be evaluated
8 individually. It would be advantageous to exploit redundant expressions contained
9 in multiple queries to significantly reduce overhead expenses associated with
10 query processing.

11 12 **SUMMARY**

13 At least one implementation described herein relates to evaluating an input
14 against multiple queries by evaluating common query expressions of the queries at
15 the same time, i.e. in parallel. An opcode block is derived from the compilation of
16 a query into a series a sequential set of instructions, or opcodes. One query
17 corresponds to one opcode block and each opcode contained there is an opcode
18 object. A logical opcode tree structure having opcode object nodes and branch
19 nodes is maintained whereby each tree segment included in a shared path
20 represents an opcode block prefix that is common to two or more queries.

21 Common opcode block prefixes are represented in the opcode tree by a single set
22 of opcode nodes so that a single execution of the opcode tree against an input
23 results in evaluation of each of the multiple queries against the input. As a result,
24 opcode objects are executed only once even though they may correspond to
25 multiple queries.

1 Branch nodes connect one tree segment to multiple dependent opcode
2 nodes. In at least one implementation, in an instance where several opcode nodes
3 that depend from a branch node perform a similar function, such as a literal
4 comparison, the branch node includes logic to perform an optimized lookup
5 routine to process the dependent opcode nodes. The optimized lookup routine
6 may include a hash routine or the like.

7 8 **BRIEF DESCRIPTION OF THE DRAWINGS**

9 A more complete understanding of exemplary systems and methods
10 described herein may be had by reference to the following detailed description
11 when taken in conjunction with the accompanying drawings wherein:

12 Fig. 1 is a block diagram of an exemplary messaging system having a filter
13 engine.

14 Fig. 2 is a block diagram of an exemplary inverse query engine.

15 Fig. 3 is a depiction of an exemplary opcode tree.

16 Fig. 4 is a flow diagram depicting a methodological implementation of
17 parallel evaluation of multiple queries.

18 Fig. 5 is a flow diagram depicting a methodological implementation of
19 adding a query to an inverse query engine that utilizes an opcode tree.

20 Fig. 6 is a diagram of an exemplary computing environment in which the
21 implementations described herein may operate.

DETAILED DESCRIPTION

The present disclosure relates to systems and methods that streamline query processing by executing similar expressions included in different queries simultaneously to reduce a number of operations that must be performed to evaluate each individual query. The examples shown and described herein specifically relate to messaging systems with inverse query engines that process multiple XPath queries against an input message.

Although specific examples are used herein, the specific examples are not intended to limit the scope of the claims appended hereto, but are provided to show one way in which the claimed subject matter may be implemented. The techniques described herein may be applied to other types of query processing systems, such as subscription services, security applications, and the like, and they may be used with other query languages.

An inverse query engine tests a collection of queries against an input and selects queries that are satisfied, or matched, by the input. Queries may also be referred to as filters, and an inverse query engine may also be referred to as a filter engine.

A basic inverse query engine separately tests an input against each of multiple queries. Such an engine becomes increasingly inefficient as the number of queries it processes grows. Filter tables (composed of queries) used in message routing systems, subscription systems and the like can contain hundreds and thousands of queries. Using a basic inverse query engine with such a system can hinder query processing within the system.

In many inverse query processing systems - e.g. messaging systems - queries included in a system filter table may be similar or have logical elements in

1 common. The techniques described herein exploit these similarities to introduce
2 parallelism into query evaluation, combining processing of redundant query
3 elements so that such an element need only be performed a single time.

4 In one or more examples shown and described herein, an inverse query
5 engine evaluates an entire filter table by executing an opcode tree that corresponds
6 to the filter table. Each query in the filter table is compiled into a linear list - i.e. a
7 sequence - of opcodes called an opcode block. A query is evaluated by executing
8 the opcode block that results from compiling the query. An opcode block of one
9 query may have a prefix in common with one or more opcode blocks of other
10 queries.

11 The opcode tree includes a root node, opcode nodes and branch nodes.
12 Each query is represented in the opcode tree by a path beginning from the root
13 node and extended through the tree. When multiple queries include a similar
14 opcode block prefix, the paths corresponding to the queries share a common
15 segment beginning with the root node until a dissimilar opcode object in the
16 respective opcode blocks is reached. At such a point in the opcode tree, a branch
17 node connects the common path segment with a diverging path segment for each
18 dissimilar opcode object that follows the similar opcode block prefix. Diverging
19 path segments can be common to multiple queries and can diverge further.

20 When an input is received for evaluation against the multiple queries, the
21 inverse query engine executes the entire opcode tree at once. As a result, opcode
22 objects that correspond to multiple queries are executed only one time and the
23 result of the execution is used in the evaluation of the input against each query that
24 corresponds to the common opcode objects.

1 Branch nodes are used to connect opcode tree segments. A branch node is
2 located in an opcode tree between a tree segment that is common to multiple
3 queries and tree branches that diverge from the common tree segment, each
4 representing a unique portion of a query opcode block. A branch node includes
5 logic for evaluating one or more branches that diverge from the branch node. The
6 logic may simply reference dependent opcode nodes, or the logic may include one
7 or more optimization techniques for processing subordinate tree segments, such as
8 using a hash table in an instance where a number of subordinate opcode objects
9 are similar literal comparisons.

10 In at least one implementation, a branch node also includes logic that
11 implements an optimization technique when it is warranted and removes the
12 optimization technique when it no longer needed. For example, a branch node
13 may be configured to implement a hash table when it detects that a certain number
14 of subordinate opcode nodes execute literal comparisons. The branch node may
15 also be configured to revert from the hash table to a linear comparison in the event
16 that the number of literal comparison opcode objects is reduced to a point where a
17 hash table is no longer effective.

18 To execute an opcode tree, an inverse query engine begins at a root node of
19 the tree and executes subordinate nodes until a branch node is reached. Each
20 branch diverging from the branch node is executed independently. For each
21 branch taken, a shallow, optimized copy of the execution context is produced. The
22 execution context contains all of the intermediate operation state gathered by
23 evaluating opcode objects. For example, if the query processor is a stack machine,
24 then the execution context would include a stack that contains results of the
25 opcode node evaluation until the branch node was reached. The assumption is that

1 performing a fast shallow copy of the context (usually just a memcpy) is more
2 efficient than independently reproducing the intermediate data multiple times –
3 and hence the rationale behind merging opcodes from multiple queries into a
4 single opcode tree. If each query is evaluated separately, the context would have
5 to be reproduced. This is a resource expense that is avoided by merging the
6 opcodes into an opcode tree. When a branch node is encountered while processing
7 a particular tree branch, the same execution process is applied recursively.

8 Intermediate Results

9 As described above, intermediate results obtained during execution of the
10 opcode tree are cached and re-used to economize processing overhead. Any
11 particular implementation of a technique for caching and retrieving intermediate
12 results may be utilized without affecting the scopes of the appended claims.

13 Regardless of the particular implementation, the execution context contains
14 all of the intermediate operation state gathered by evaluating opcode nodes.
15 During the execution of the opcode tree, when a branch node or a leaf node is
16 reached, the execution context has the result of the execution to that point.

17 When a branch node is reached and individual branches therefrom are
18 executed, the manner in which the execution context is handled depends on a
19 particular way in which a query processor operates. One implementation utilizes a
20 stack machine query processor. In a stack machine, an execution context includes
21 a stack in which intermediate operands and opcode node evaluation results are
22 stored. Opcodes push and pop data and nodesets to and from the stack.

23 For each branch taken in the present case, an optimized copy of the
24 evaluation context is made. This is a simple procedure (typically just a memcpy)
25 that is more efficient than reproducing the evaluation context multiple times (once

1 for each branch taken). In a typical situation, if each query was evaluated
2 independently, redundant opcodes would have to be re-executed and the execution
3 context reproduced multiple times.

4 In another implementation, a variable machine could be utilized instead of a
5 stack machine. When using a query language that has no side effect (such as
6 XPath), the query processor could utilize variables, where each unique
7 intermediate result corresponds to a unique variable.

8 In this implementation, the evaluation context includes a variable table.
9 Each opcode node in the opcode tree that produces a unique result (e.g., an opcode
10 that adds two numbers produces a new result) is associated with a predefined
11 variable. An opcode node evaluates by setting its associated variable in the
12 evaluation context.

13 When a branch node is encountered, the execution context does not need to
14 be reproduced since each opcode node is associated with a unique variable, i.e. a
15 variable that, once set, doesn't change during the execution of the opcode tree.
16 While this implementation is more efficient in some terms, one problem associated
17 with it is that the variable table can be quite large, thereby making some aspects of
18 the present invention more difficult to implement.

19 Therefore, a particular implementation for handling intermediate results
20 must take into account certain characteristics of an execution environment.
21 However, any method known in the art for caching and retrieving intermediate
22 results may be utilized with the presently described systems and methods.

23 Opcode Tree Copy

24 In some systems, it is important that an opcode tree be available for
25 maintenance (adding and/or removing nodes) at the same time that it can be used

1 to process queries. In one implementation a copy of the opcode tree may also be
2 maintained so that opcodes can be merged into or removed from the opcode tree
3 while the opcode tree copy can be contemporaneously used for query processing.
4 In an alternative implementation, a read/write precedence may be established
5 wherein read operations related to the opcode tree have precedence over write
6 operations. As a result, the opcode tree can be used for processing during an
7 operation to merge or remove opcodes from the tree since the opcode tree since no
8 opcodes can be added to the opcode tree during a merging process.

9 These and other features are described in greater detail, below, with respect
10 to the figures which depict specific exemplary implementations.

11 **Exemplary Computer System**

12 Fig. 1 is a simplified block diagram of an exemplary computer system 100
13 in accordance with the presently described systems and methods. The exemplary
14 computer system 100 as shown encompasses at least a portion of a messaging
15 system 114 in memory 102. The computer system 100 also includes a processor
16 for executing computer instructions and an input/output (I/O) module 106 for
17 receiving and transmitting data to/from the computer system 100. The computer
18 system 100 also includes a number of hardware components 108 that are not
19 specified herein, but are necessary to support operation of the computer system
20 100.

21 The memory 102 stores an operating system 110 that controls operation of
22 the computer system 100 and one or more software applications 112 that support
23 particular functions of the computer system 100. The messaging system 114
24 includes a filter engine 116 and a filter table 118 which are used to process
25 messages.

1 The computer system 100 is configured to communicate over a network
2 118 with a remote computing device 120 that includes a messaging system
3 component 122. The messaging system component 122 is configured to send
4 messages to the messaging system 114 of the computer system 100 over the
5 network 118, which may be a local area network (LAN), a wide area network
6 (WAN), the Internet or the like.

7 **Exemplary Inverse Query Engine**

8 **Fig. 2** is a block diagram of an inverse query engine 200 in accordance with
9 the present description. The inverse query engine 200 is similar to the filter engine
10 116 shown in Fig. 1. In the following discussion, reference may be made to
11 particular elements and reference numerals shown in Fig. 1.

12 The inverse query engine 200 includes an input module 202 that is
13 configured to receive queries 204 and messages 206. Although the input module
14 202 is depicted as a single logical unit, in practice separate input modules may be
15 utilized. In such a configuration, a first input module would be configured to
16 receive input queries and a second input module would be configured to receive
17 input messages.

18 In the present example, the queries are XPath queries and the messages are
19 constructed according to a protocol compatible therewith, such as the Simple
20 Object Access Protocol (SOAP). It is noted, however, that the queries and the
21 messages may be formatted according to one or more other languages and/or
22 protocols.

23 The inverse query engine 200 also includes a query processor 208 that is
24 configured to execute instructions within the inverse query engine 200 to provide
25 the functionality thereto, such as processing input messages against queries, etc. A

1 language analysis module 210 is configured to recognize and handle queries input
2 into the inverse query engine 200. The inverse query engine 200 also includes at
3 least one filter table 212 that stores multiple queries (query_A 214, query_B 216
4 through query_n 218). The filter table 212 includes a reference 219 to an opcode
5 tree 230.

6 The language analysis module 210 includes components that recognize and
7 compile queries that are added to the inverse query engine 200. In the present
8 example, the language analysis module 210 includes a lexical analysis module
9 (“lexer”) 220, a parser 222 and a compiler 224. The lexer 220 receives an input
10 and groups characters included in the input into elemental language units (tokens)
11 which are then sent to the parser 222. The parser 222 analyzes the tokens received
12 from the lexer 220 and breaks them into constituent parts. The compiler 224
13 translates the information from the lexer 220 and the parser 222 into opcodes that
14 the query processor 208 uses to process the input queries.

15 The compiler 224 generates opcodes 226, from each query that it compiles.
16 Each query produces a set of opcode objects called an opcode block. An opcode
17 merger 228 merges opcode blocks containing newly compiled opcode objects into
18 the opcode tree 230.

19 The opcode tree 230 includes opcode nodes objects and branch nodes as
20 described above. Each opcode node comprises an opcode object. The opcode tree
21 230 is a logical structure that includes opcode nodes, some of which include
22 references to other opcode nodes to create a linked list. Creation and maintenance
23 of the opcode tree 230 will be discussed in greater detail below.

24 The inverse query engine 200 also maintains an opcode tree copy 232 that
25 is used during processing so that operations may be undertaken on the opcode tree

230 without interfering with query processing. Updates are made to the opcode tree copy 232 while the opcode tree 230 remains available for query processing. When the updates are complete, the updated opcode tree copy 232 is stored as the opcode tree 230. Thus, a current version of the opcode tree 230 is always available for query processing.

An intermediate results cache 234 is include in the inverse query engine 200 to store intermediate results obtained during execution of the opcode tree, and execution contexts related to opcode tree processing. The contents of the intermediate results cache 234 are used as described above to avoid re-execution of opcode nodes that are common to more than one query.

The elements shown in the inverse query engine 200 will be discussed in greater detail below, with respect to the following figures.

Exemplary Opcode Tree

Fig. 3 is a depiction of an exemplary opcode tree 300 having opcode nodes and branch nodes. For discussion purposes, an exemplary filter table that includes the queries listed below in **Table 1** is presented. It is noted that the queries shown in Table 1 do not conform to exact XPath query syntax. Rather, the queries have been simplified to better demonstrate the structure and function of the opcode tree 300. For example, the XPath query expression

`/env:Envelope/env:Headers/wsa:Action = "http://abc.com"`

is simply represented below (as Query 1) as “action = http://abc.com”. The highlighted reference letters shown in brackets (“[]”) adjacent to certain query expressions identify the corresponding segments in the opcode tree 300.

Q1: action = http://abc.com [a]
Q2: action = http://abc.com [a] AND weight = 32 [b]
Q3: action = http://xyz.com [c] AND (weight = 40 [d] OR priority = 1 [e])
Q4: action = http://abc.com [a] AND weight = 32 [b] AND priority > 1 [f]
Q5: priority = 1 [g]
Q6: action = http://xyz.com [c]
Q7: (action = http://xyz.com [c] AND weight < 40 [h]) OR priority > 1 [i]

Table 1: Exemplary Filter Table Queries

Using the segment identifiers [a] - [i], the queries can be expressed briefly as follows:

Q1 = a
Q2 = a, b
Q3 = c, d, e
Q4 = a, b, f
Q5 = g
Q6 = c
Q7 = c, h, i

In the above example, the query compiler (a simplistic one) has compiled Q2 into an opcode block containing 2 opcodes: a, b. Opcode 'a' will evaluate the expression "action = http://abc.com". The opcodes are also labeled 'a' and 'b' for illustration— to show correspondence to expressions. It is apparent that opcode blocks for some queries contain opcodes that are contained in other queries. For example, Q1, Q2 and Q4 all begin with opcode [a]. It is this type of similarity between queries that is exploited in the use of the opcode tree.

The opcode tree 300 is structured in accordance with the group of queries shown in Table 1. The opcode tree 300 has a root node 302 and a first branch node 304. The first branch node 304 is the root node of multiple sub-trees, or branches which are executed recursively. Each branch is traversed sequentially

1 and opcode nodes are executed as they are encountered. The traversal and
2 execution are described in greater detail below, with respect to Fig. 5.

3 Opcode node 306 of the opcode tree 300 is associated with expression **[a]**
4 of Query 1 (Table 1). Since Query 1 consists of a simple expression (i.e., it is not
5 a compound expression made up of multiple sub-expressions), the expression
6 contained in Query 1 - "action = "http://abc.com" - is compiled and the resulting
7 opcode object is associated with opcode node 306. Traversing the opcode tree 300
8 with regard to an input from the root node 302 through the first branch node 304 to
9 opcode node 306 executes Query 1 (Q1).

10 Since Query 2 and Query 4 also begin with expression **[a]**, opcode node
11 306 can also be used in the evaluation of these queries. Thus, two additional
12 evaluations of expression **[a]** are avoided by using the result associated with
13 opcode node 306.

14 The second opcode segment of Query 2 - expression **[b]** - follows
15 immediately after expression **[a]**. Therefore, opcode node 308 is created as a child
16 node of opcode node 306 to represent expression **[b]**. Query 2 is represented by
17 the opcode tree path of root node 302, first branch node 304, opcode node 306 and
18 opcode node 308.

19 Query 4 includes expression **[a]** and expression **[b]** but also includes
20 expression **[f]**. Therefore, opcode node 310, which corresponds to expression **[f]**,
21 is created as child node of node 308. Query 4 is represented by an opcode tree
22 path of root node 302, first branch node 304, opcode node 306, opcode node 308
23 and opcode node 310.

24 There are no other queries that begin with expression **[a]**, so a new branch
25 is created from the first branch node 304 to correspond with a different expression

1 that begins a query. In the present example, Query 5 begins with expression [g],
2 so opcode node 312 is created off of the first branch node 304. Since no query
3 contains any expressions subsequent to expression [g], opcode node 312 does not
4 have any child nodes.

5 There are other queries that begin with an expression other than the
6 expressions that have already been mapped (expressions [a], [b], [f] and [g]), so a
7 new branch is created off of the first branch node 304 for these other expressions.
8 Opcode node 314 is created off of the first branch node 304 to represent
9 expression [c], which begins Query 3, Query 6 and Query 7.

10 It is noted that the queries and the opcode tree 300 are both of a hierarchical
11 nature. The opcode tree 300 has different levels that correspond to opcode
12 segments in different positions. The first level in the opcode tree 300 under the
13 root node corresponds with an initial opcode segment of the queries. In the
14 present example, there are three distinct initial opcode segments in the queries that
15 correspond to three distinct nodes in the first level of the tree.

16 Since there are no distinct opcode segments that start a query other than
17 those represented by opcode node 306, opcode node 312 and opcode node 314,
18 there are no additional child nodes subordinate to the first branch node 304.

19 Query 3 and Query 7 begin with the same query segment ([c]) that is
20 represented by opcode node 314 but include subsequent segments. Therefore,
21 Query 3 and Query 7 will require a separate branch off of opcode node 314. A
22 second branch node 316 initiates the branches for Query 3 and Query 7. Query 3
23 includes the subsequent expressions [d] and [e], so these expressions make up the
24 branch associated with Query 3. Opcode node 318 corresponds to expression [d]
25 and opcode node 320 corresponds to expression [e].

1 A branch corresponding to Query 7 is created off of the second branch node
2 316. Query 7 includes two expressions after expression [c], namely expression [h]
3 and expression [i]. Therefore, the branch includes opcode node 322 that
4 corresponds to expression [h] and opcode node 324 that corresponds to expression
5 [i].

6 **Exemplary Methodological Implementation: Parallel Query Evaluation**

7 Fig. 4 is a flow diagram 400 depicting a methodological implementation of
8 parallel evaluation of multiple queries using the opcode tree techniques described
9 herein. In the following discussion of Fig. 4, continuing reference will be made to
10 the elements and reference numerals included in Figures 2 and 3.

11 At block 402, the message 206 is received via the input module 202 of the
12 inverse query engine 200. The message 206 is input to the query processor 208
13 for evaluation against the filter table 212. To perform the evaluation, the nodes of
14 the opcode tree 300 are traversed and each opcode node is executed. Initially,
15 control is directed to the root node 302 (block 404) of the opcode tree 300. In the
16 present example, the root node 302 is not associated with any executable query
17 opcode segment but in some instances, the root node 302 may contain opcodes
18 that will be executed at block 404. If an opcode object is executed for the root
19 node, the intermediate result is cached at block 406. In most instances, block 406
20 will not be necessary.

21 If there is a next node to evaluate ("Yes" branch, block 408) it is
22 determined if the next node is a branch node at block 410. At this point in the
23 present example, the next node is the first branch node 304. If the next node is a
24 branch node ("Yes" branch, block 410), the process proceeds with a following
25 node of a first branch off the first branch node at block 408. If the node is not a

1 branch node ("No" branch, block 410), then the opcode node is executed at block
2 414 and the intermediate result is cached at block 416 (although this step may not
3 be necessary in most instances).

4 The process then reverts to block 408 and subsequent nodes are processed
5 until a leaf node at the end of a branch is reached. At that point, there are no more
6 nodes to be processed in the currently executed branch ("No" branch, block 408).
7 If there is a previous branch node ("Yes" branch, block 418), then focus is returned
8 to the previous branch node at block 420. If there is another branch from that
9 branch node ("Yes" branch, block 422), then the process reverts to block 408 for
10 additional processing of nodes in the next branch off the branch node.

11 If there is not another branch from the branch node ("No" branch, block
12 422), then it is determined if there is a previous branch node at block 418. If so,
13 the process repeats as described above. If there is not a previous branch node
14 ("No" branch, block 418), then control is at the highest branch node in the opcode
15 tree 300 and there are no more branches from that branch node that need to be
16 evaluated. Therefore, the opcode tree 300 has been executed in its entirety and the
17 message 206 has been evaluated against each query in the filter table 212.

18 Queries that were satisfied by the input message 206 are returned at block
19 424. In the previous example of the message evaluated against the queries
20 included in Table 1, the matching queries are Query 1 and Query 7. The inverse
21 query engine 200 then performs post-processing on the matching queries, such as
22 taking an action associated with one or more of the matching queries.

23
24
25

Exemplary Methodological Implementation: Merging Opcodes

Fig. 5 is a flow chart 500 depicting a methodological implementation for merging a query into an opcode tree. In the following discussion, continuing reference will be made to the elements and reference numerals of Figures 2 and 3.

At block 502, the input module 202 receives a query 204 to add to the filter table 224. The query 204 is directed to the language analysis module 210 where it is parsed and compiled into opcode objects 226 (block 504). The opcode objects 226 are then ready to be merged into the opcode tree 230.

In the following discussion, the opcode merger 232 handles the process of merging the opcode objects 226 into the opcode tree 230. However, in another implementation, opcode objects are configured to merge themselves into the opcode tree 230. In such a case, the logic required to understand the relationship between an opcode object and the opcode tree is embedded in an instance of an object that is created when a query is compiled. The latter implementation allows new optimization techniques to be added without having to implement logic into the opcode merger. Only an object would have to be updated. This implementation is described in greater detail below.

The opcode merger 232 takes an available opcode object ("Yes" branch block 506) and determines if an identical opcode object is already in the opcode tree 230, beginning the analysis at the root node 302. If the opcode object is already in the opcode tree ("Yes" branch, block 508), then the process focuses on that node in the tree and a next opcode object - if available - is processed beginning at block 506.

1 If the opcode object is not in the opcode tree (“No” branch, block 508),
2 then it must be determined if a branch node exists immediately below the last node
3 that corresponded to a query opcode object already in the opcode tree 300. For
4 example, referring to the opcode tree 300 in Fig. 3, suppose that a new query being
5 merged into the tree was structured as “Action = xyz.com/123”. Since “xyz.com”
6 is already represented by opcode node 314 (expression [c]), a new branch would
7 have to be created for the “/123” portion of the query. Opcode node 314 already
8 has two branches so there is an existing branch node 316 (“Yes” branch, block
9 512). Therefore, a new opcode node will be added as a branch off of the existing
10 branch node 316. If there was not an existing branch node (“No” branch, block
11 512), a new branch node would be created at block 514.

12 Before a new opcode node is added to the opcode tree 300, the opcode
13 merger 228 evaluates the context of the branch node to determine if an
14 optimization can be made to the branch node (block 516). If the branch node can
15 be optimized (“Yes” branch, block 518), the branch node is optimized and
16 implemented at block 520. The new opcode node is then added as a new branch
17 off of the branch node (block 522).

18 If the branch node cannot be optimized (“No” branch, block 518), the new
19 opcode node is simply added as a new node off of the branch node at block 522.
20 Optimization techniques are described in greater detail below.

21 The process is repeated from block 506 until there are no more opcode
22 nodes to merge into the opcode tree 230. If the previously described operations
23 have been performed on an opcode tree copy 232 to preserve the opcode tree 230
24 for processing during the merge operation, the opcode tree copy 232 is stored as
25 the opcode tree 230 at this time.

Optimizations

There are situations where an optimization technique can be implemented to increase the efficiency of the query evaluation process. In some systems, a majority of expressions included in queries perform similar operations. For example, queries written in XPath frequently perform literal comparisons. Table 2 includes some exemplary XPath expressions that could be included in one or more queries.

(1) /Envelope/Headers/Action = 'http://abc.com/xyz'
(2) /Envelope/Headers/Action = 'http://abc.com/xyz'
and /Envelope/Headers/Weight = 32
(3) /Envelope/Headers/Hopcount/TTL = 54
(4) /XYZ/Priority > 3 and /XYZ/Priority <=5

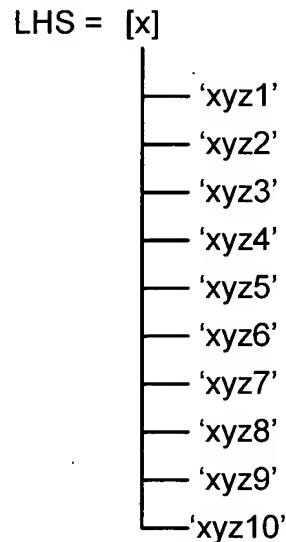
Table 2: Exemplary XPath Queries

All of the queries shown in Table 2 test literals against source values and return a Boolean result (T or F). Such queries are compiled into expressions using opcodes that are designed to perform literal comparisons.

For example, suppose the following Xpath expressions are to be tested against a given input:

LHS = 'xyz1'
LHS = 'xyz2'
LHS = 'xyz3'
LHS = 'xyz4'
LHS = 'xyz5'
LHS = 'xyz6'
LHS = 'xyz7'
LHS = 'xyz8'
LHS = 'xyz9'
LHS = 'xyz10'

1 The opcode merger is configured to combine opcode blocks derived from
2 compiling the expressions into an opcode tree logically arranged as follows, where
3 [x] represents a branch node:



12 The prefix common to each expression ("LHS =") is blended into the
13 opcode tree as an opcode node, and each diverging literal forms a different branch
14 from the opcode node. When an input message is processed against an opcode
15 tree containing this structure, the value associated with "LHS" is retrieved once.
16 However, to compute the Boolean result of each branch, the query processor 208
17 performs a literal comparison for each branch. As a general rule, if there were n
18 such expressions, the query processor would have to perform n string comparisons
19 to find the result for each literal comparison.

20 In the exemplary case, it would be much more efficient to utilize a hash
21 function to optimize the comparisons. When merging opcode objects into an
22 opcode tree, the opcode merger 228 is configured to detect a situation in which an
23 optimization algorithm can be implemented and combines literal comparisons into
24

1 an indexed literal branch opcode object. The specialized branch opcode takes the
2 place of a generic branch node in the opcode tree.

3 An indexed literal branch opcode includes a branch index that contains
4 (key, value) pairs. The literals form the key, and an identifier associated with a
5 branch of a literal is associated with its value. When a query containing the literal
6 is evaluated, the value of the literal is hashed to derive an index entry that
7 references the literal value. In some cases, multiple entries will be identified.

8 Branch index objects allow the result of every branch to be obtained very
9 quickly, frequently by making a single comparison. In the above example, the
10 value of "LHS" is retrieved, hashed and looked up in the index. The index returns
11 a set of (key, value) pairs, where the key is the value of matching literals and the
12 value identifies branches to which the literal belongs. Every branch that is
13 identified evaluates to True. The other branches, therefore, evaluate to False.

14 Indexed branch opcode objects provide a general framework for insertion
15 of indexing techniques that are customized to the type of comparison involved.
16 Types of comparisons that may be implemented with indexed branch opcode
17 literals include, but are not limited to, string and number equality literals using
18 hash tables, interval trees to optimize numeric interval queries, opcodes that
19 perform string prefix matching using tries (a standard computer science data
20 structure designed for optimal string prefix matching), and the like.

21 Optimizing opcode objects may sometimes require devolving from a
22 specialized optimization technique to a linear comparison. In the example of a
23 hash table, it may only be more efficient to utilize a hash table when there are, say,
24 more than fifty entries in the hash table. If such a case, removing a few opcode
25

1 nodes may make the hashing operation more expensive than a literal comparison
2 procedure.

3 The opcode merger 228 is configured to recognize such situations and
4 remove an optimization technique if the removal will increase efficiency. In an
5 implementation where an opcode merger 228 is not used, opcode objects will
6 include logic to recognize these particular situations and to reform an indexed
7 branch literal opcode object, for example, into a generic branch opcode object.

8 Removing opcode nodes from the opcode tree are described in more detail
9 below.

10 **Removing Opcode Nodes**

11 Sometimes it is necessary to remove opcode nodes from an opcode tree
12 230. In such a situation, the inverse query engine 200 is requested to remove a
13 particular query. To remove opcode nodes associated with the query, a component
14 of the inverse query engine 200 (such as the opcode merger 228 or the query
15 processor 208) traverses the opcode tree 230 and attempts to identify tree
16 segments that are common between the query being removed and at least one
17 other query in the opcode tree.

18 The opcode tree is traversed according to the query to be removed until a
19 branch node is encountered that begins a branch specific to the query to be
20 removed. Only the branch that is specific to the query is removed. In cases where
21 removal of the specific branch causes removal of one of two branches, a branch
22 node that was the point of divergence for the two branches is removed, since a
23 branch node at that location is no longer required.

Exemplary Computer Environment

The various components and functionality described herein are implemented with a computing system. Fig. 6 shows components of typical example of such a computing system, i.e. a computer, referred by to reference numeral 600. The components shown in Fig. 6 are only examples, and are not intended to suggest any limitation as to the scope of the functionality of the invention. Furthermore, the invention is not necessarily dependent on the features shown in Fig. 6.

Generally, various different general purpose or special purpose computing system configurations can be used. Examples of well known computing systems, environments, and/or configurations that may be suitable for use with the invention include, but are not limited to, personal computers, server computers, hand-held or laptop devices, multiprocessor systems, microprocessor-based systems, set top boxes, programmable consumer electronics, network PCs, minicomputers, mainframe computers, distributed computing environments that include any of the above systems or devices, and the like.

The functionality of the computers is embodied in many cases by computer-executable instructions, such as program modules, that are executed by the computers. Generally, program modules include routines, programs, objects, components, data structures, etc. that perform particular tasks or implement particular abstract data types. Tasks might also be performed by remote processing devices that are linked through a communications network. In a distributed computing environment, program modules may be located in both local and remote computer storage media.

1 The instructions and/or program modules are stored at different times in the
2 various computer-readable media that are either part of the computer or that can be
3 read by the computer. Programs are typically distributed, for example, on floppy
4 disks, CD-ROMs, DVD, or some form of communication media such as a
5 modulated signal. From there, they are installed or loaded into the secondary
6 memory of a computer. At execution, they are loaded at least partially into the
7 computer's primary electronic memory. The invention described herein includes
8 these and other various types of computer-readable media when such media
9 contain instructions programs, and/or modules for implementing the steps
10 described below in conjunction with a microprocessor or other data processors.
11 The invention also includes the computer itself when programmed according to
12 the methods and techniques described below.

13 For purposes of illustration, programs and other executable program
14 components such as the operating system are illustrated herein as discrete blocks,
15 although it is recognized that such programs and components reside at various
16 times in different storage components of the computer, and are executed by the
17 data processor(s) of the computer.

18 With reference to Fig. 6, the components of computer 600 may include, but
19 are not limited to, a processing unit 602, a system memory 604, and a system bus
20 606 that couples various system components including the system memory to the
21 processing unit 602. The system bus 606 may be any of several types of bus
22 structures including a memory bus or memory controller, a peripheral bus, and a
23 local bus using any of a variety of bus architectures. By way of example, and not
24 limitation, such architectures include Industry Standard Architecture (ISA) bus,
25 Micro Channel Architecture (MCA) bus, Enhanced ISA (EISAA) bus, Video

1 Electronics Standards Association (VESA) local bus, and Peripheral Component
2 Interconnect (PCI) bus also known as the Mezzanine bus.

3 Computer 600 typically includes a variety of computer-readable media.
4 Computer-readable media can be any available media that can be accessed by
5 computer 600 and includes both volatile and nonvolatile media, removable and
6 non-removable media. By way of example, and not limitation, computer-readable
7 media may comprise computer storage media and communication media.
8 “Computer storage media” includes volatile and nonvolatile, removable and non-
9 removable media implemented in any method or technology for storage of
10 information such as computer-readable instructions, data structures, program
11 modules, or other data. Computer storage media includes, but is not limited to,
12 RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM,
13 digital versatile disks (DVD) or other optical disk storage, magnetic cassettes,
14 magnetic tape, magnetic disk storage or other magnetic storage devices, or any
15 other medium which can be used to store the desired information and which can be
16 accessed by computer 600. Communication media typically embodies computer-
17 readable instructions, data structures, program modules or other data in a
18 modulated data signal such as a carrier wave or other transport mechanism and
19 includes any information delivery media. The term “modulated data signal”
20 means a signal that has one or more of its characteristics set or changed in such a
21 manner as to encode information in the signal. By way of example, and not
22 limitation, communication media includes wired media such as a wired network or
23 direct-wired connection and wireless media such as acoustic, RF, infrared and
24 other wireless media. Combinations of any of the above should also be included
25 within the scope of computer readable media.

1 The system memory 604 includes computer storage media in the form of
2 volatile and/or nonvolatile memory such as read only memory (ROM) 608 and
3 random access memory (RAM) 610. A basic input/output system 612 (BIOS),
4 containing the basic routines that help to transfer information between elements
5 within computer 600, such as during start-up, is typically stored in ROM 608.
6 RAM 610 typically contains data and/or program modules that are immediately
7 accessible to and/or presently being operated on by processing unit 602. By way
8 of example, and not limitation, Fig. 6 illustrates operating system 614, application
9 programs 616, other program modules 618, and program data 620.

10 The computer 600 may also include other removable/non-removable,
11 volatile/nonvolatile computer storage media. By way of example only, Fig. 6
12 illustrates a hard disk drive 622 that reads from or writes to non-removable,
13 nonvolatile magnetic media, a magnetic disk drive 624 that reads from or writes to
14 a removable, nonvolatile magnetic disk 626, and an optical disk drive 628 that
15 reads from or writes to a removable, nonvolatile optical disk 630 such as a CD
16 ROM or other optical media. Other removable/non-removable,
17 volatile/nonvolatile computer storage media that can be used in the exemplary
18 operating environment include, but are not limited to, magnetic tape cassettes,
19 flash memory cards, digital versatile disks, digital video tape, solid state RAM,
20 solid state ROM, and the like. The hard disk drive 622 is typically connected to
21 the system bus 606 through a non-removable memory interface such as data media
22 interface 632, and magnetic disk drive 624 and optical disk drive 628 are typically
23 connected to the system bus 606 by a removable memory interface such as
24 interface 634.
25

1 The drives and their associated computer storage media discussed above
2 and illustrated in Fig. 6 provide storage of computer-readable instructions, data
3 structures, program modules, and other data for computer 600. In Fig. 6, for
4 example, hard disk drive 622 is illustrated as storing operating system 615,
5 application programs 617, other program modules 619, and program data 621.
6 Note that these components can either be the same as or different from operating
7 system 614, application programs 616, other program modules 618, and program
8 data 620. Operating system 615, application programs 617, other program
9 modules 619, and program data 621 are given different numbers here to illustrate
10 that, at a minimum, they are different copies. A user may enter commands and
11 information into the computer 600 through input devices such as a keyboard 636
12 and pointing device 638, commonly referred to as a mouse, trackball, or touch
13 pad. Other input devices (not shown) may include a microphone, joystick, game
14 pad, satellite dish, scanner, or the like. These and other input devices are often
15 connected to the processing unit 602 through an input/output (I/O) interface 640
16 that is coupled to the system bus, but may be connected by other interface and bus
17 structures, such as a parallel port, game port, or a universal serial bus (USB). A
18 monitor 642 or other type of display device is also connected to the system bus
19 606 via an interface, such as a video adapter 644. In addition to the monitor 642,
20 computers may also include other peripheral output devices 646 (e.g., speakers)
21 and one or more printers 648, which may be connected through the I/O interface
22 640.

23 The computer may operate in a networked environment using logical
24 connections to one or more remote computers, such as a remote computing device
25 650. The remote computing device 650 may be a personal computer, a server, a

1 router, a network PC, a peer device or other common network node, and typically
2 includes many or all of the elements described above relative to computer 600.

3 The logical connections depicted in Fig. 6 include a local area network (LAN) 652
4 and a wide area network (WAN) 654. Although the WAN 654 shown in Fig. 6 is
5 the Internet, the WAN 654 may also include other networks. Such networking
6 environments are commonplace in offices, enterprise-wide computer networks,
7 intranets, and the like.

8 When used in a LAN networking environment, the computer 600 is
9 connected to the LAN 652 through a network interface or adapter 656. When used
10 in a WAN networking environment, the computer 600 typically includes a modem
11 658 or other means for establishing communications over the Internet 654. The
12 modem 658, which may be internal or external, may be connected to the system
13 bus 606 via the I/O interface 640, or other appropriate mechanism. In a networked
14 environment, program modules depicted relative to the computer 600, or portions
15 thereof, may be stored in the remote computing device 650. By way of example,
16 and not limitation, Fig. 6 illustrates remote application programs 660 as residing
17 on remote computing device 650. It will be appreciated that the network
18 connections shown are exemplary and other means of establishing a
19 communications link between the computers may be used.

1 **Conclusion**

2 Although details of specific implementations and embodiments are
3 described above, such details are intended to satisfy statutory disclosure
4 obligations rather than to limit the scope of the following claims. Thus, the
5 invention as defined by the claims is not limited to the specific features described
6 above. Rather, the invention is claimed in any of its forms or modifications that
7 fall within the proper scope of the appended claims, appropriately interpreted in
8 accordance with the doctrine of equivalents.

9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25